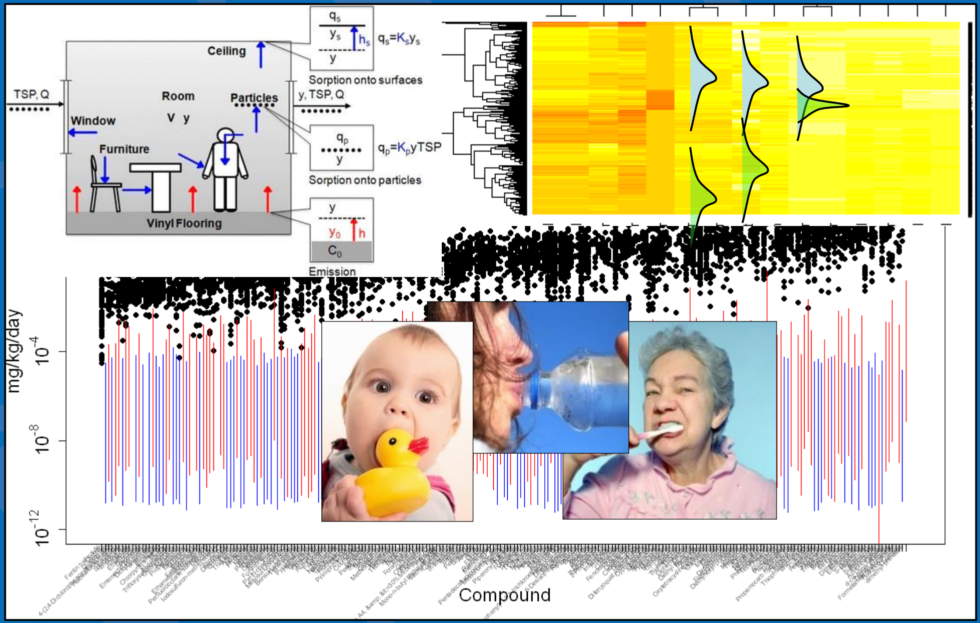
EPA Exposure Research and the ExpoCast Project: New Methods and New Data

John Wambaugh
National Center for Computational Toxicology
U.S. EPA, Office of Research and Development



NIEHS Exposure Science & the Exposome Webinar
May 14, 2015

Introduction

- The timely characterization of the human and ecological risk posed by thousands of existing and emerging commercial chemicals is a critical challenge facing EPA in its mission to protect public health and the environment
- While advances have been made in HT toxicity screening, **exposure** and **dosimetry** prediction methods applicable to 1000s of chemicals are needed

November 29, 2014

Thousands of Chemicals

- Wild (2005): The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology
  - Examination of the small molecule metabolites that constitute the metabolome offers opportunities to address exposure
  - Park *et al*. (2012): 3221 chemicals in humans, many appear to be exogenous

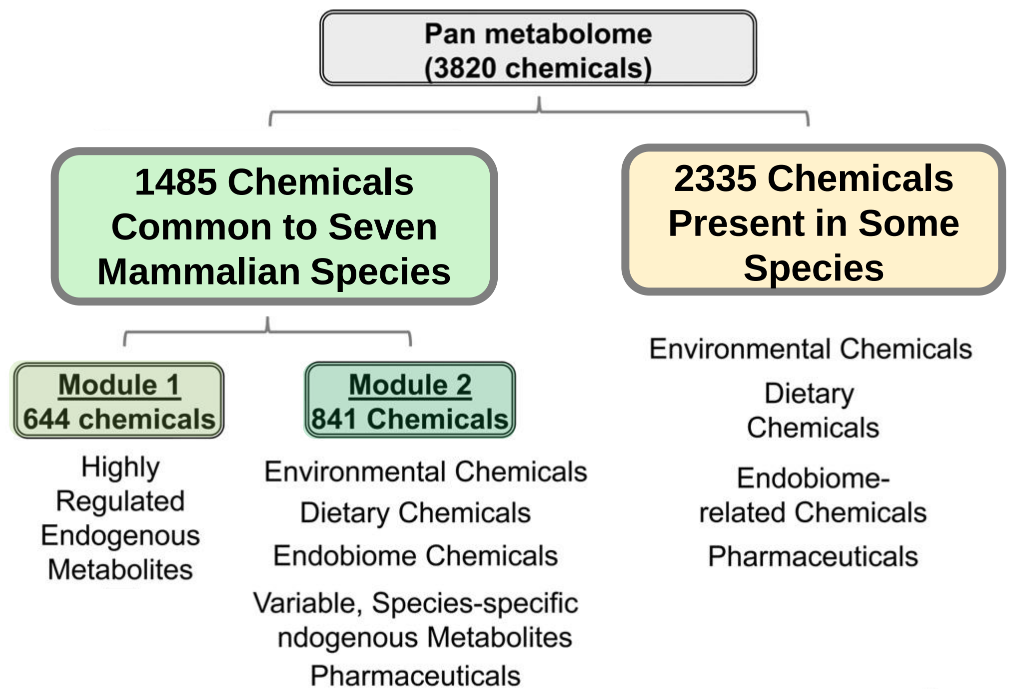
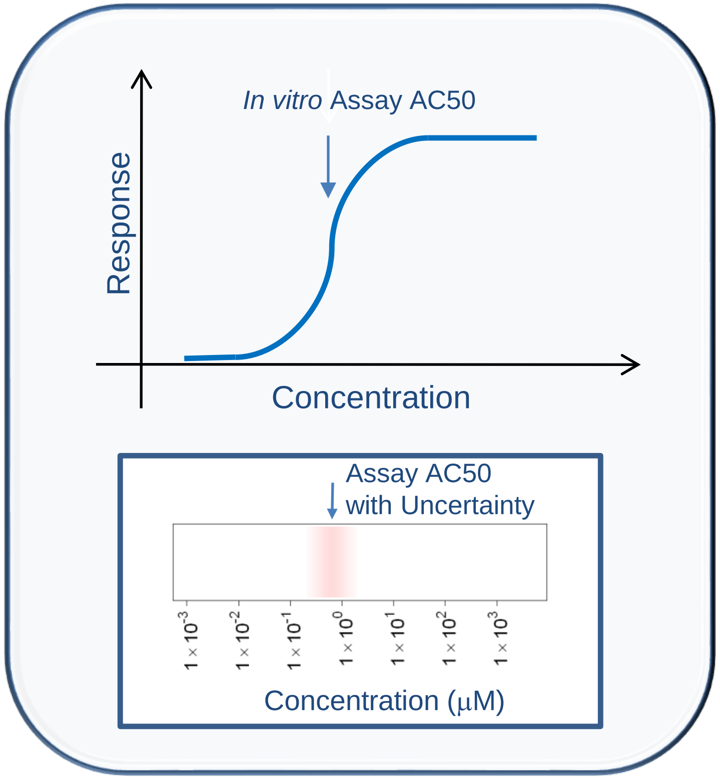


Figure modified from Park et al. (2012)

United States
Environmental Protection
Agency

High-Throughput Bioactivity

- **Tox21**: Examining >10,000 chemicals using ~50 assays intended to identify interactions with biological pathways (Schmidt, 2009)
- **ToxCast**: For a subset (>1000) of Tox21 chemicals ran >500 additional assays (Judson et al., 2010)
- Most assays conducted in dose-response format (identify 50% activity concentration – AC50 – and efficacy if data described by a Hill function)
- All data is public: http://actor.epa.gov/

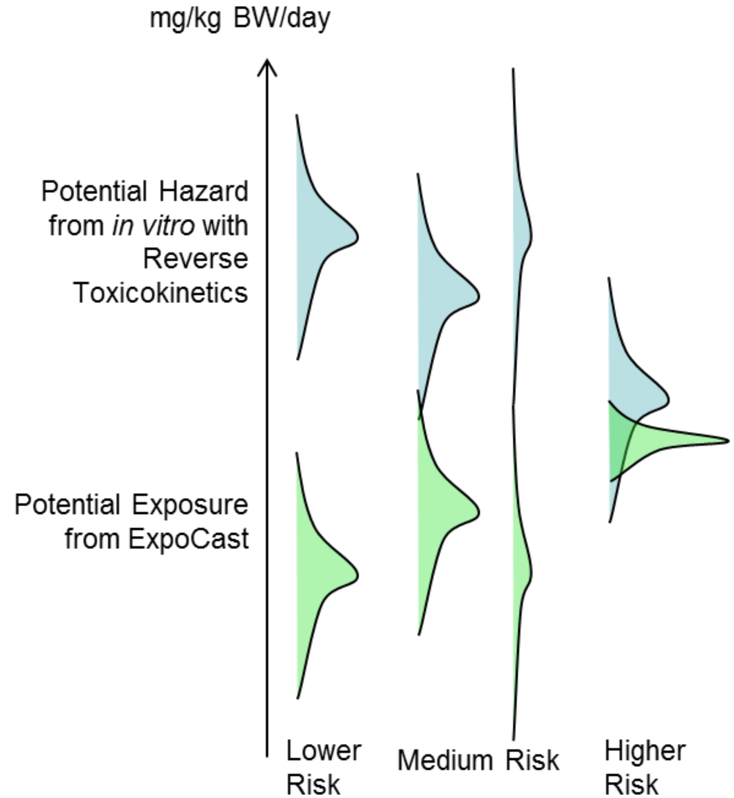


Risk Prioritization Requires Exposure

- **Tox21/ToxCast**: Examining thousands of chemicals using high throughput screening assays to identify *in vitro* concentrations that perturb biological pathways (Schmidt, 2009)

- In Wetmore *et al.* (2012), High throughput toxicokinetic *in vitro* methods are used to approximately convert *in vitro* bioactive concentrations (µM) into daily doses needed to produce similar levels in a human (mg/kg BW/day)

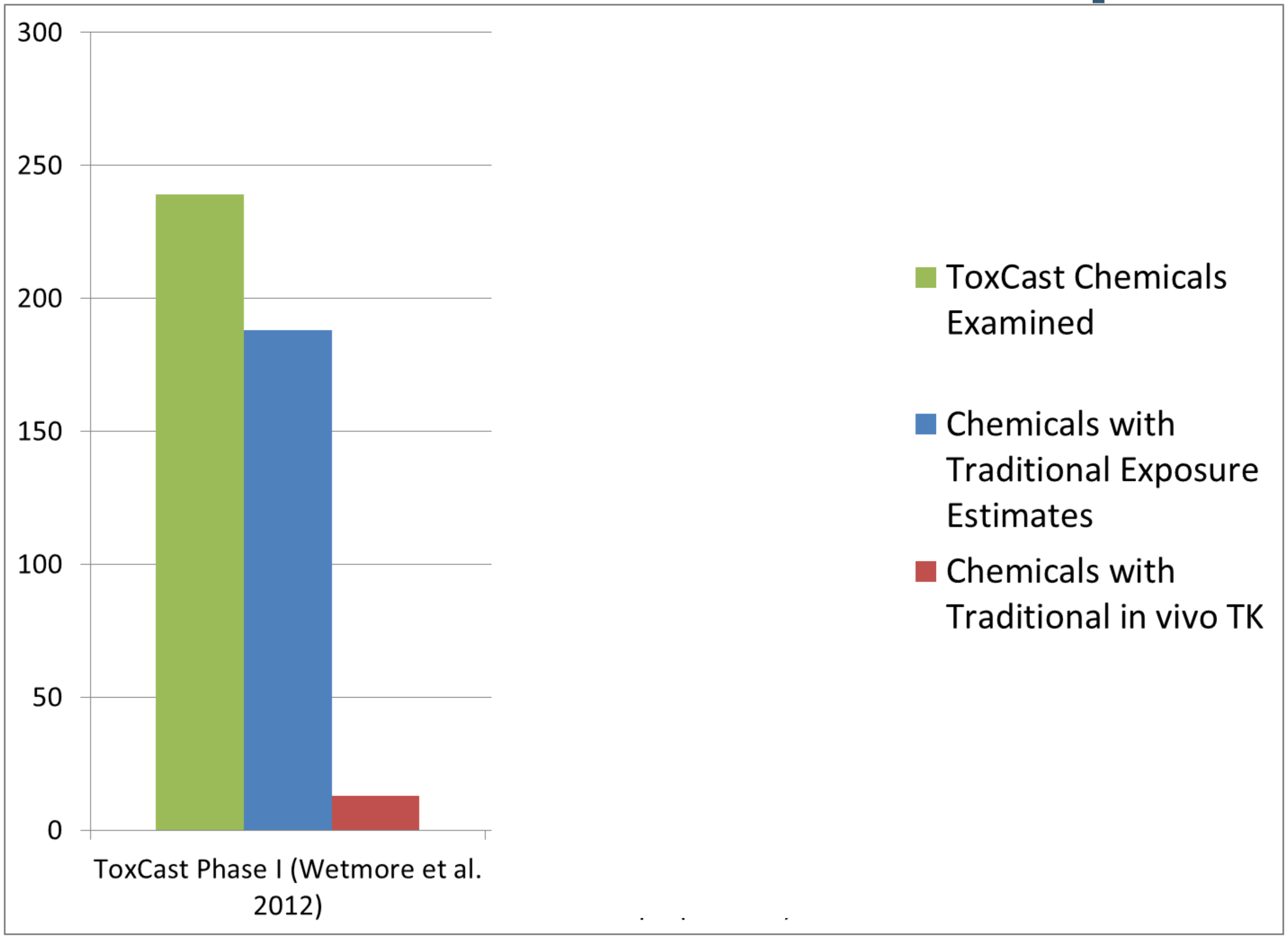
- These doses can then be directly compared with exposure rates, ***where available***



December, 2014 Panel:
"Scientific Issues Associated with Integrated Endocrine Bioactivity and Exposure-Based Prioritization and Screening"

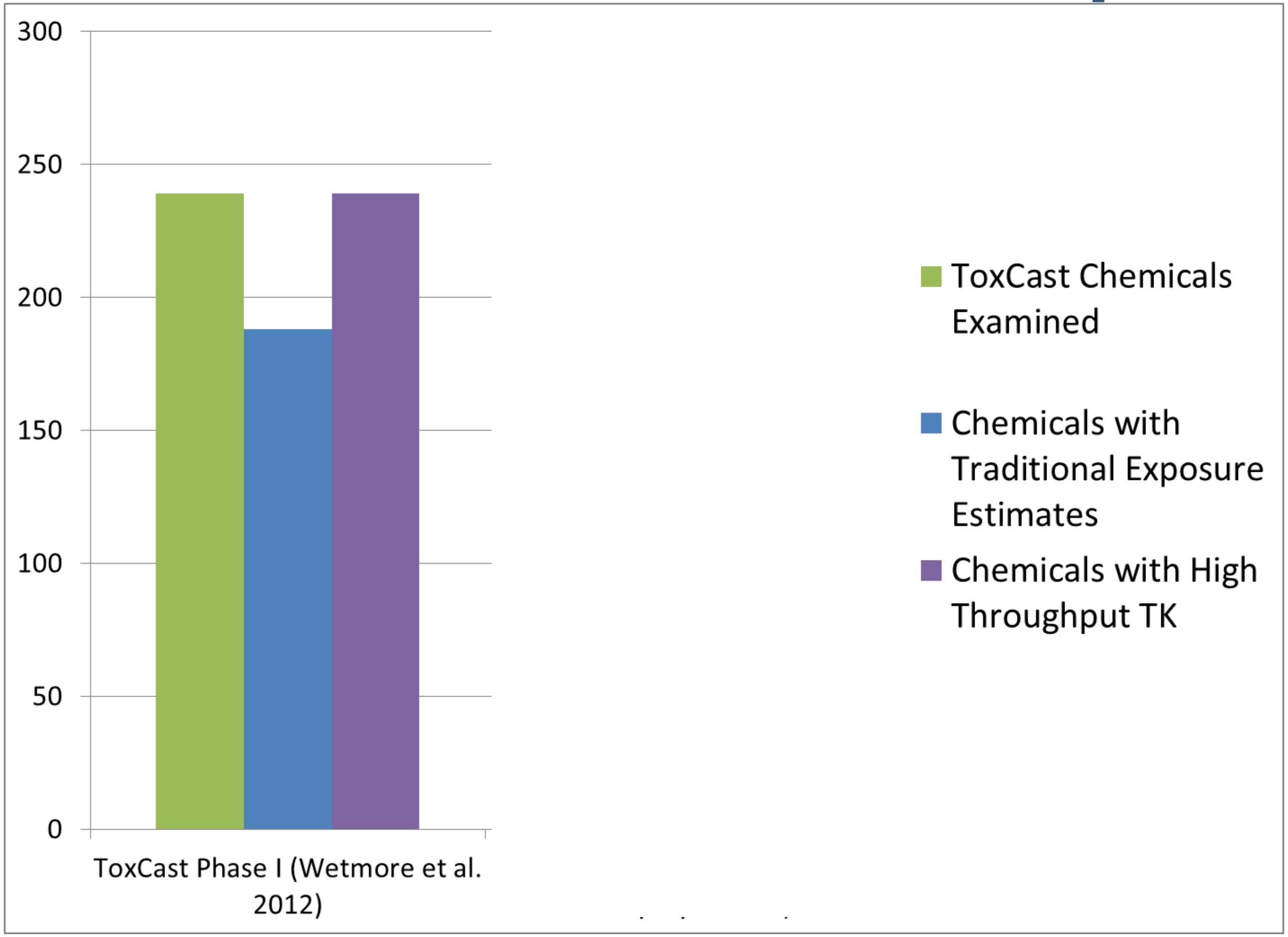
DOCKET NUMBER EPA–HQ–OPP–2014–0614

In Vitro Bioactivity, *In Vivo* Toxicokinetics, and Exposure



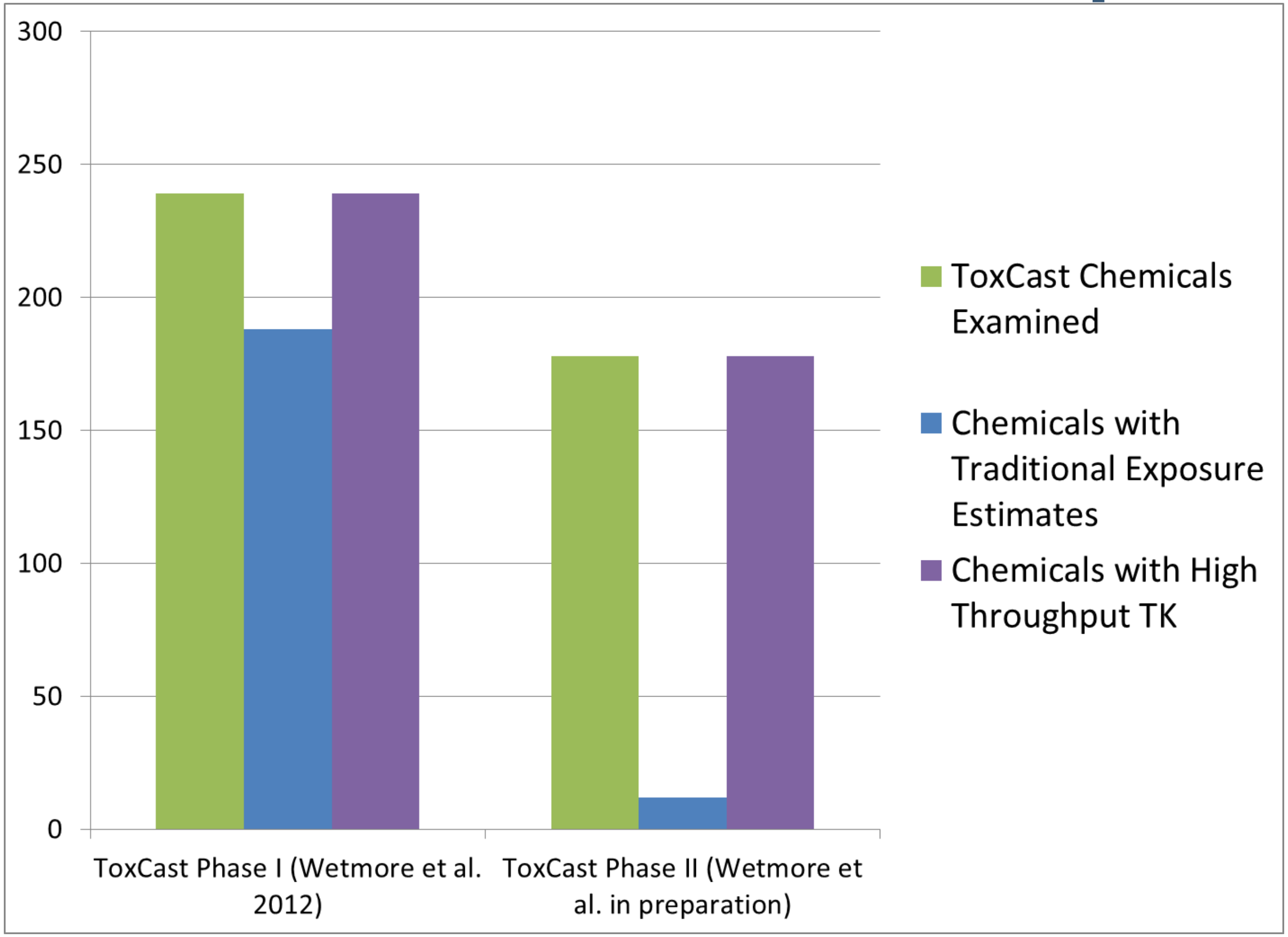
- ToxCast Phase I focused on relatively data-rich pesticides – but there was little TK data

In Vitro Bioactivity, *In Vitro* Toxicokinetics, and Exposure



- Studies like Wetmore et al. (2012),addressed the need for TK data using *in vitro* methods

In Vitro Bioactivity, *In Vitro* Toxicokinetics, and Exposure



- For non-pesticide chemical space, there is a paucity of data for providing context to HTS data (Egeghy *et al.* (2012))

Extending Network Analysis to Develop the Exposome

- Chemical exposure perturbs coupled networks spanning multiple levels of biological organization
- Measurements show not just one xenobiotic chemical, but thousands (the Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology)
- High throughput exposure tools are essential for understanding the Exposome

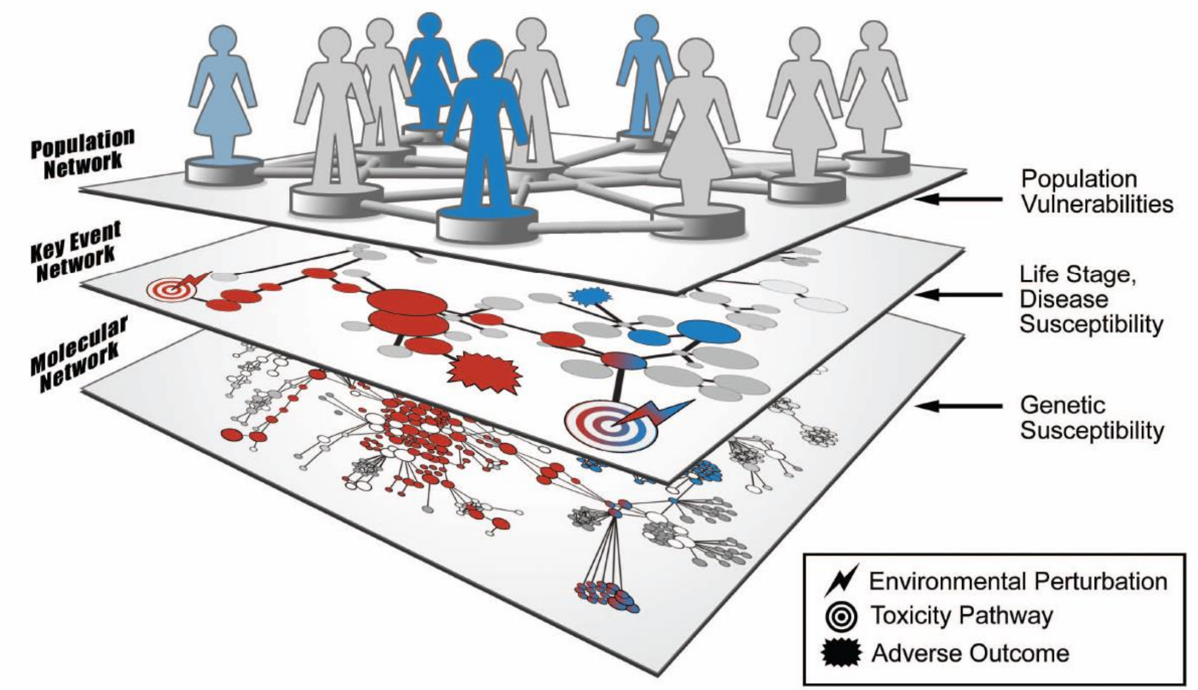


Figure modified from Edwards and Preston (2008)

Goals for High Throughput Exposure

- Exposure Forecasting: ExpoCast
- Incorporate multiple models into consensus predictions for 1000s of chemicals
- Evaluate/calibrate predictions with available measurement data across many chemical classes
- Empirically estimate uncertainty in predictions

United States Environmental Protection Agency

High Throughput Exposure Forecasts

- New methods for Exposure Forecasting (ExpoCast) currently being considered for prioritization of chemical testing in the Endocrine Disrupter Screening Program (EDSP)
- Favorably reviewed by July 2014 Federal Insecticide, Fungicide, Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP)

https://federalregister.gov/a/2014-12593

Agency/Docket Numbers:

EPA-HQ-OPP-2014-0331

FRL-9910-22

Exposure SAP White Paper

New High-throughput Methods to Estimate Chemical Exposure

Scientific Advisory Panel Meeting, July 2014

Exposure Space

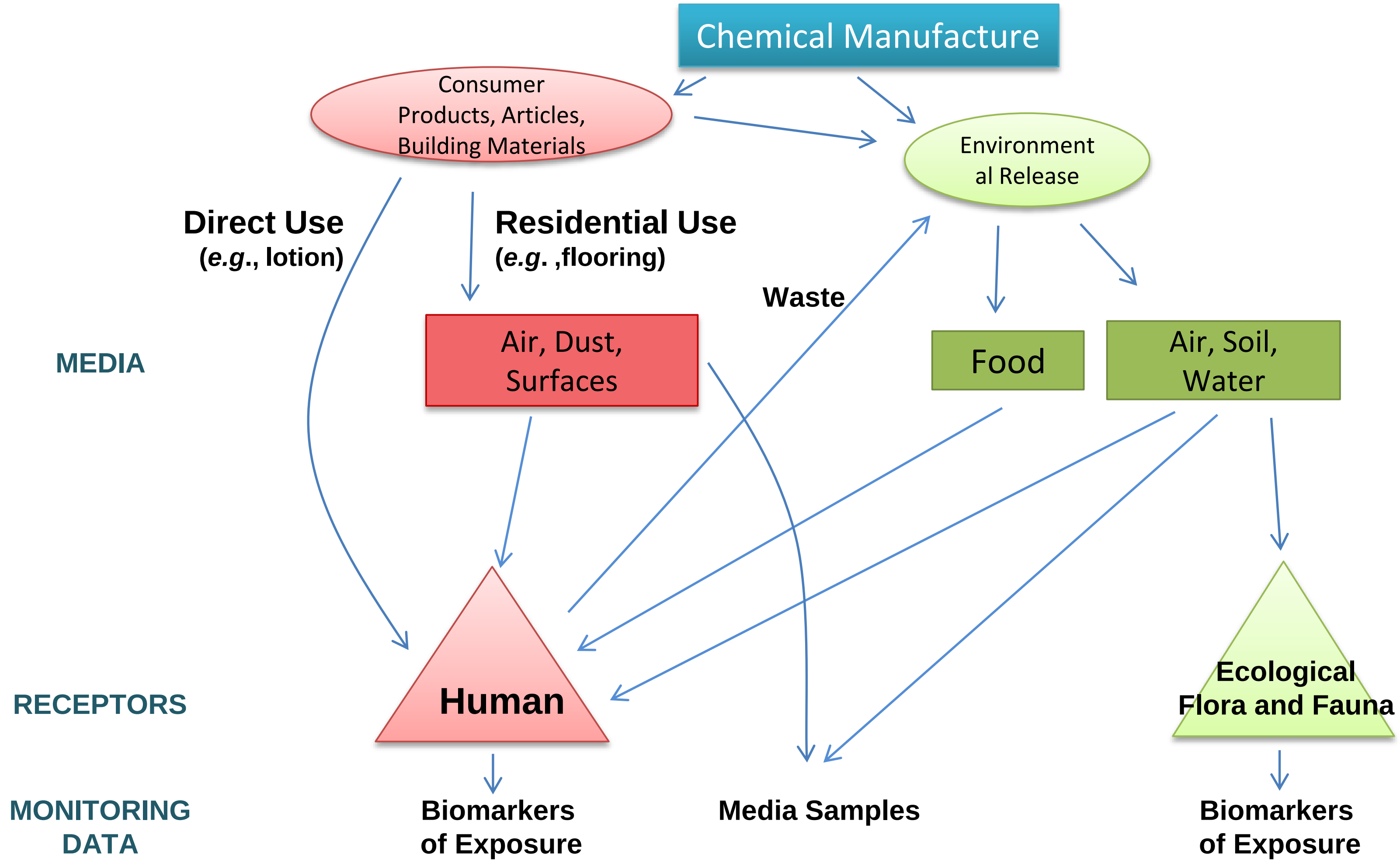


Figure from Kristin Isaacs

Exposure Pathways

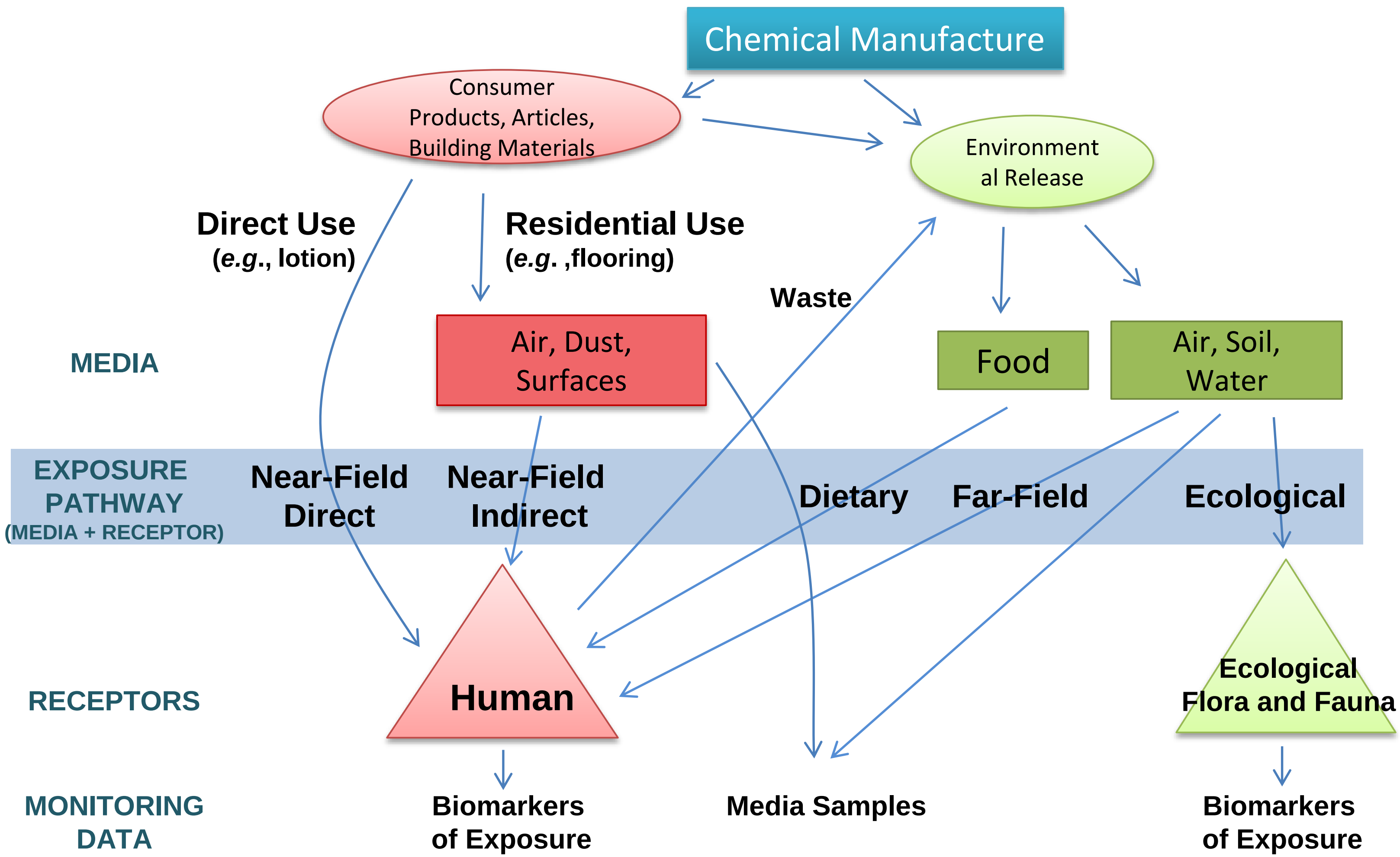


Figure from Kristin Isaacs

Forward Modeling of Exposure Pathways

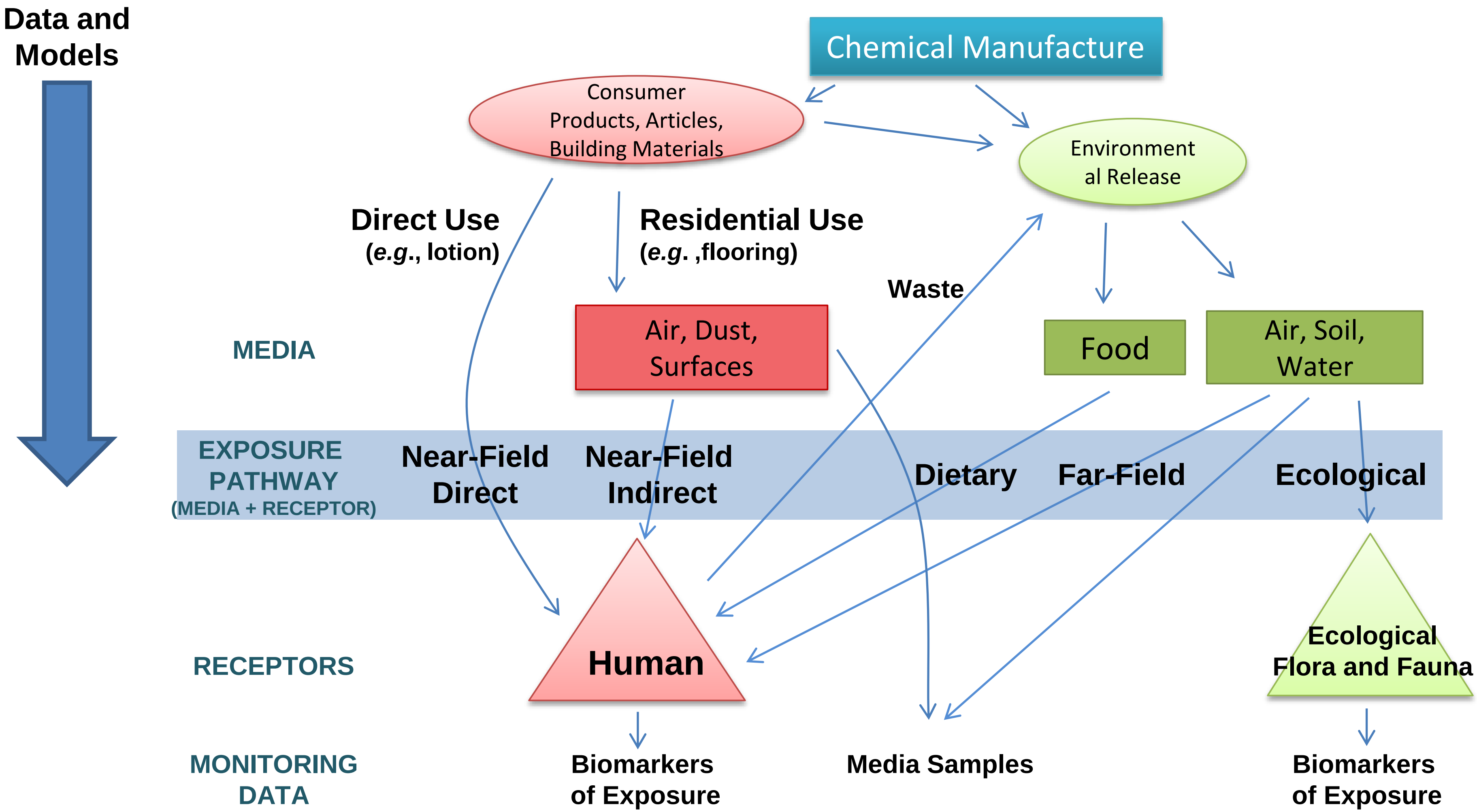


Figure from Kristin Isaacs

Inference of Exposure Pathways

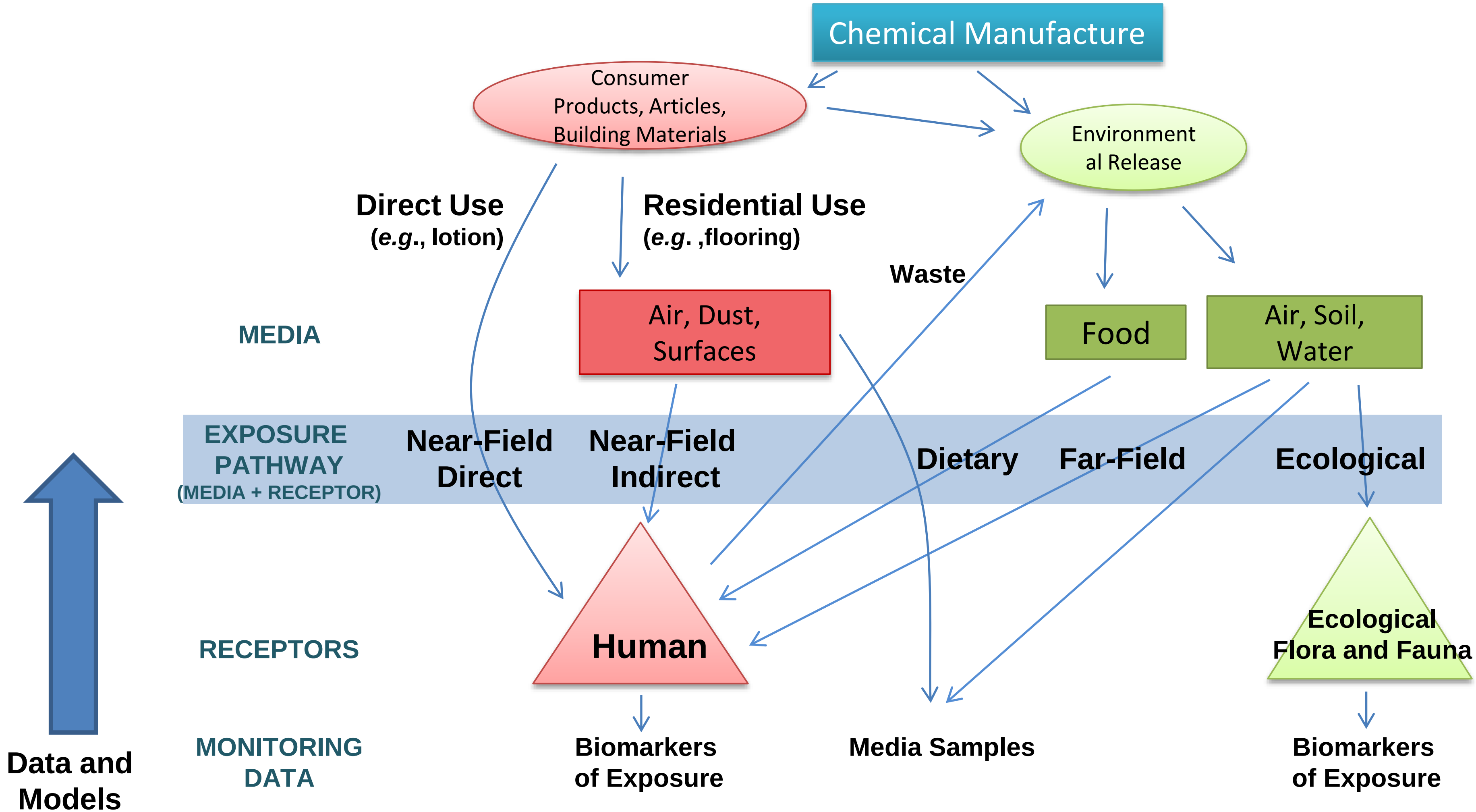


Figure from Kristin Isaacs

Evaluation of Forward Predictions with Inferred Exposure

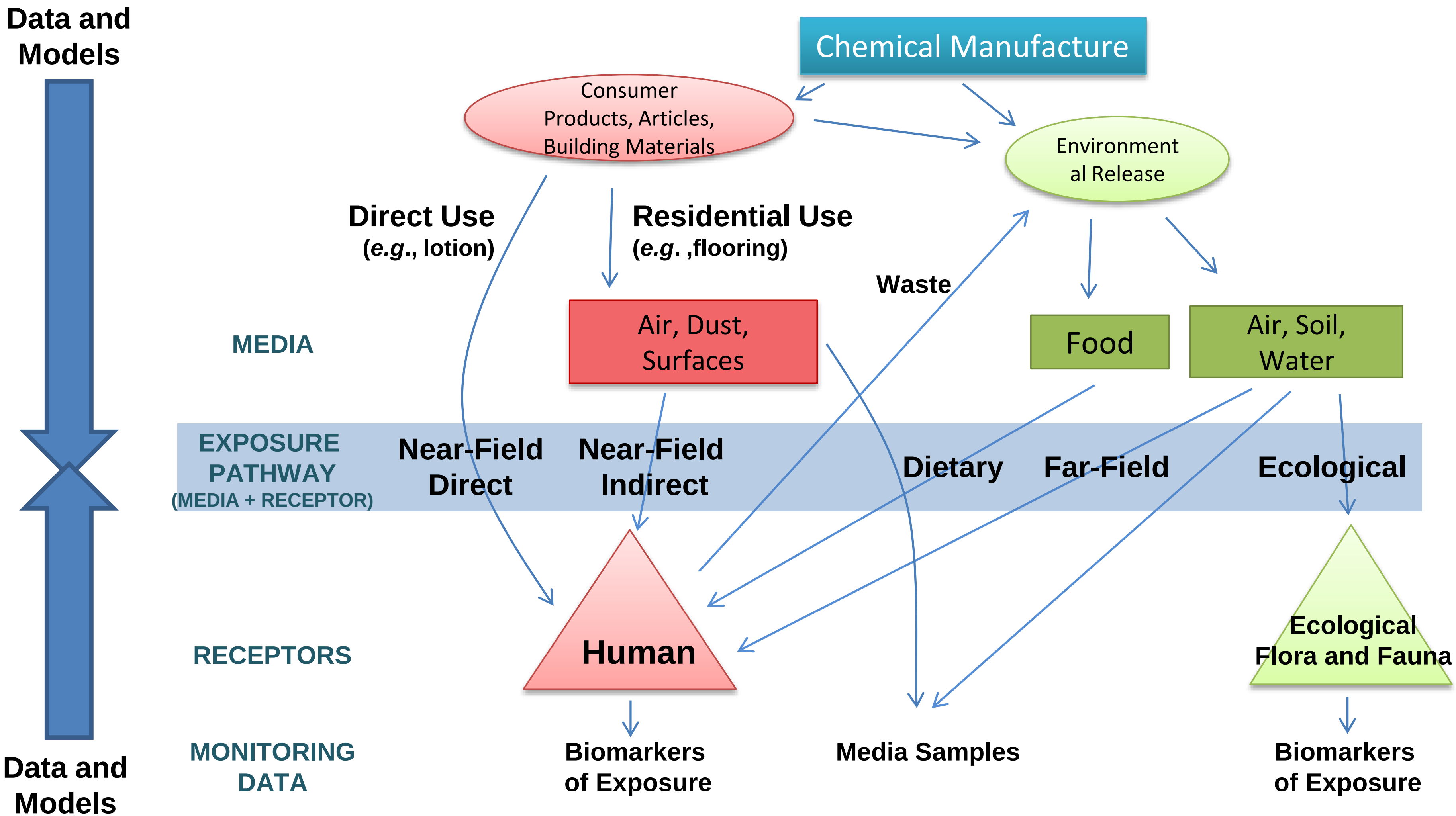
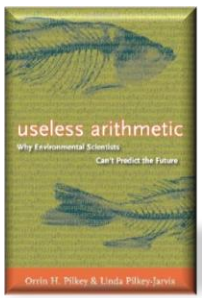


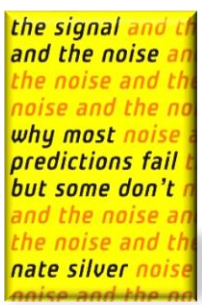
Figure from Kristin Isaacs

How To Deal with Uncertainty

“Useless Arithmetic: Why Environmental Scientists Can’t Predict the Future” (Pilkey & Pilkey-Jarvis, 2007)

“The Signal and the Noise: Why Many Predictions Fail – But Some Don’t” (Silver, 2012):



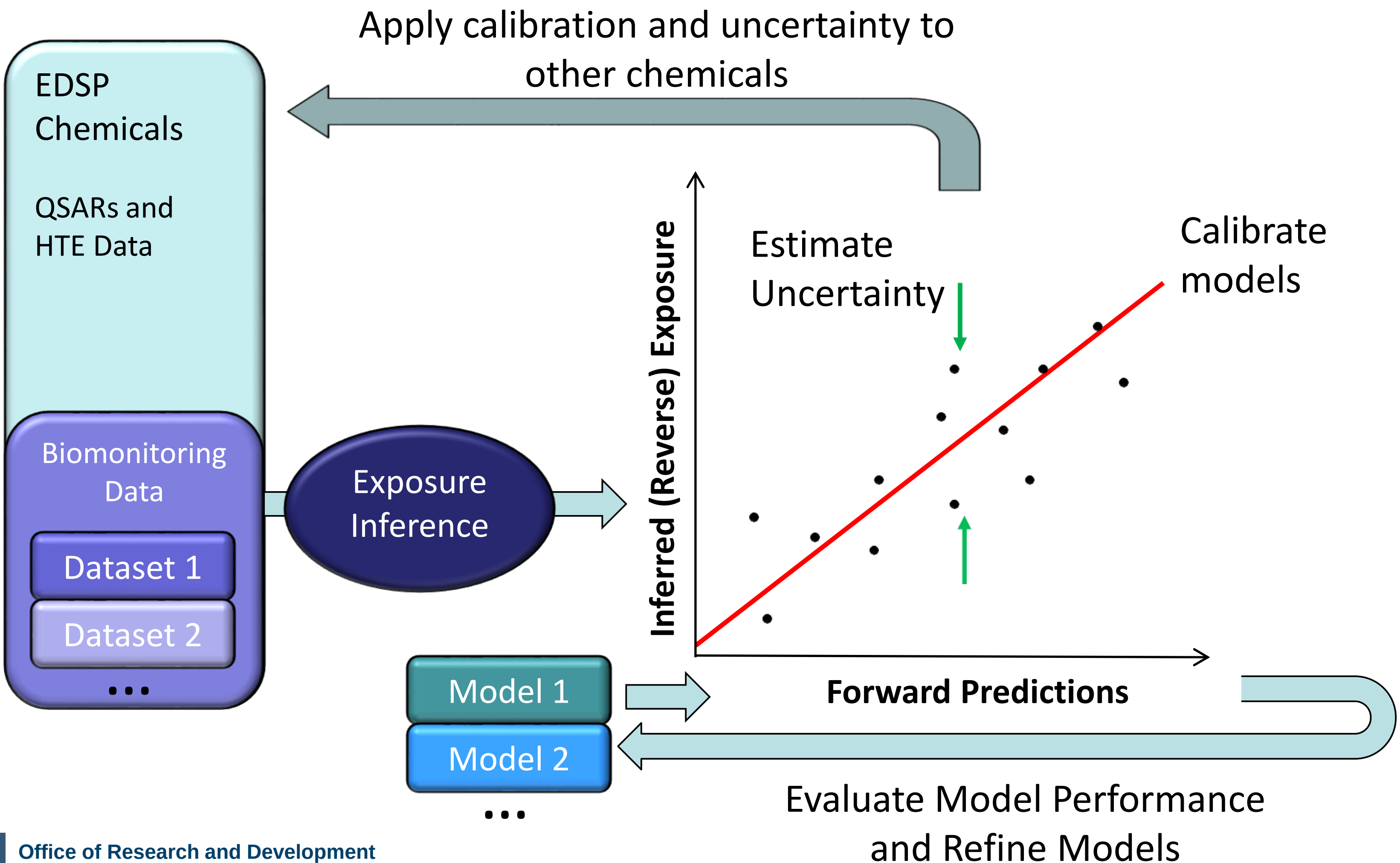


1) Think probabilistically (especially, Bayesian): Consider an approach that evaluates model performance systematically across as many chemicals (and chemistries) as possible

2) Forecasts change: Today’s forecast reflects the best available data today but we must accept that new data and new models will cause predictions to be revised

3) Look for consensus: Evaluate as many models and predictors/ predictions as possible

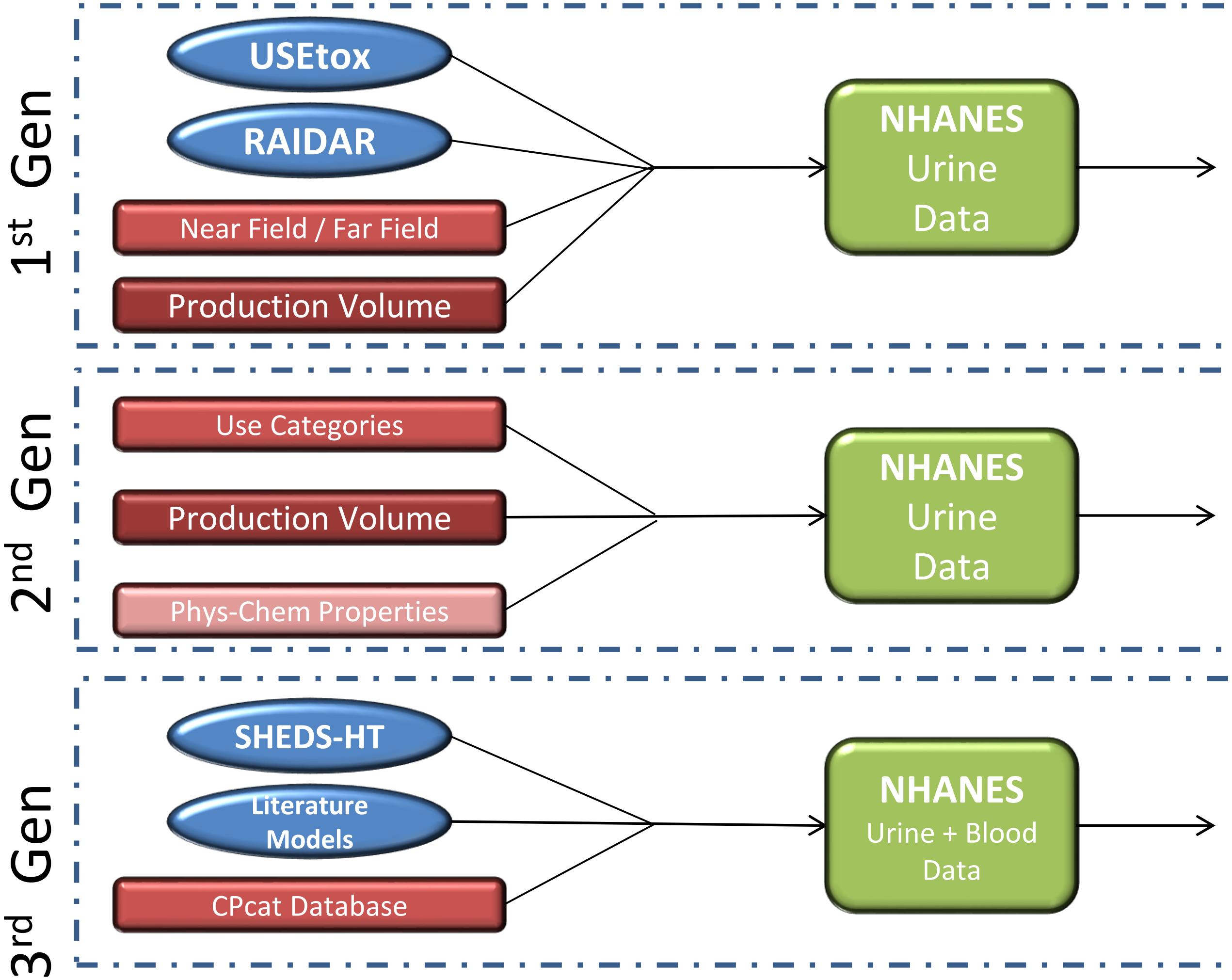
Systematic Empirical Evaluation of Models



SEEM Evolution – Human Exposure

United States
Environmental Protection
Agency

Model and Predictors | **Calibration / Evaluation Data** | **SEEM Conclusion**



1st Gen

- Existing complex fate and transport models have low correlation to measured exposures
- Near field factor most important

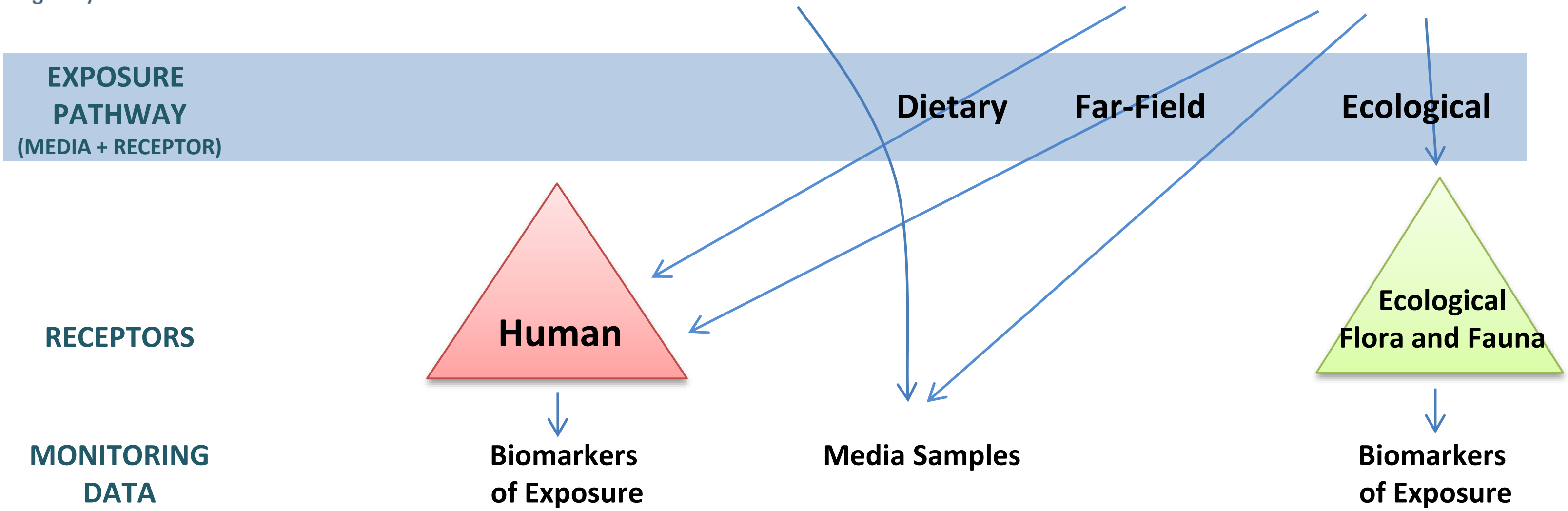
2nd Gen

- Simple, readily available data
- Better correlation to measured exposures
- Similar predictions across demographics

3rd Gen

- Analysis ongoing

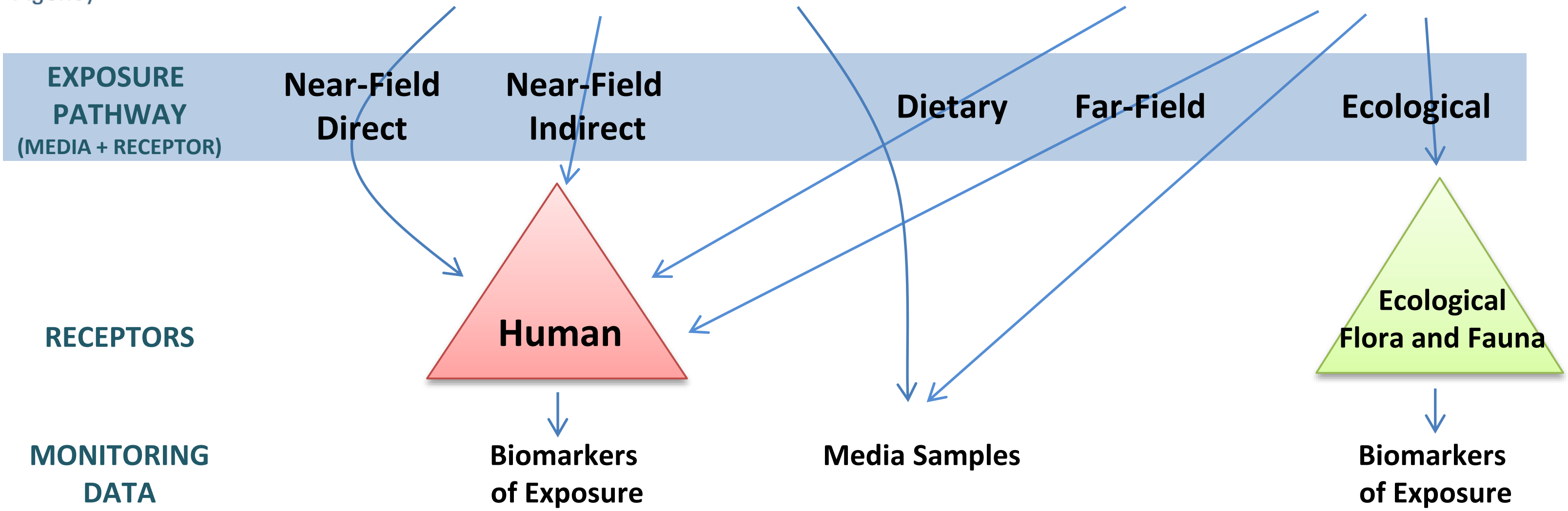
Coverage of Exposure Pathways: Data and Models



- Prior to 2013, most high throughput exposure models covered far-field (industrial release) and dietary (industry + agricultural pesticides) exposure pathways (Rosenbaum et al., 2006, Arnot et al. 2006)
- Usage of chemicals that are likely to have high exposure from these routes has been reduced (e.g., PCBs, DDT)
- Far-field chemicals and pesticide actives likely to be below NHANES limit of detection (Wambaugh et al., 2013, 2014)

Figure modified from original by Kristin Isaacs

Coverage of Exposure Pathways: Data and Models



- Far-field chemicals and pesticide actives likely to be below NHANES limit of detection, **but** chemicals with near field sources of exposure are often detectable in the urine of NHANES participants (Wambaugh et al., 2013, 2014) (TEAM study found this for volatiles – Wallace et al., 1987)

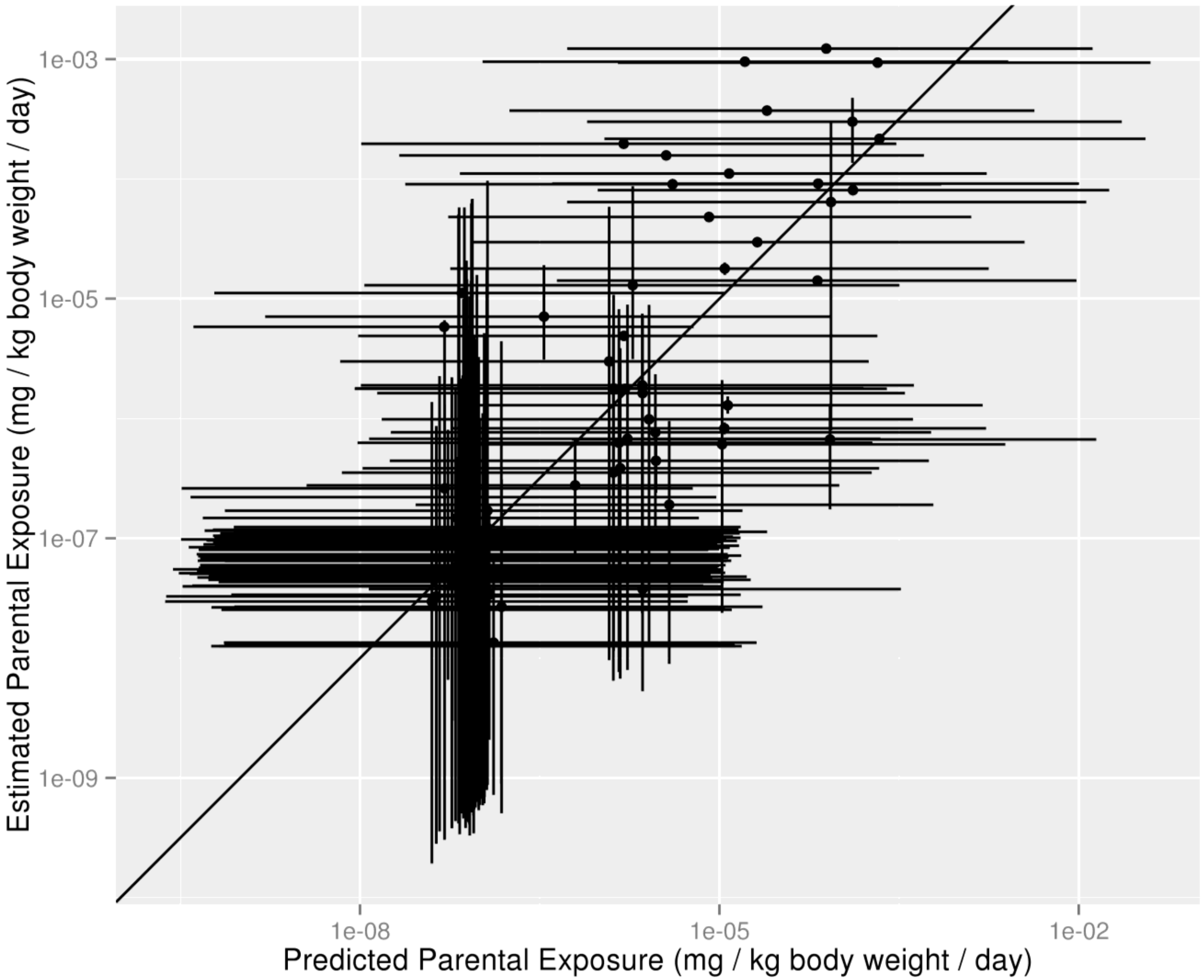
- New data sources developed to characterize chemical usage in consumer products (*e.g.*, Goldsmith et al., 2014, Dionisio et al., 2015)

- New models developed to characterize exposure to thousands of these chemicals (*e.g.*, Isaacs et al., 2014)

Figure modified from original by Kristin Isaacs

Predicting NHANES exposure rates

United States
Environmental Protection
Agency



$R^2 \approx 0.5$ indicates that we can predict 50% of the chemical to chemical variability in mean NHANES exposure rates

Same five predictors work for all NHANES demographic groups analyzed – stratified by age, sex, and body-mass index

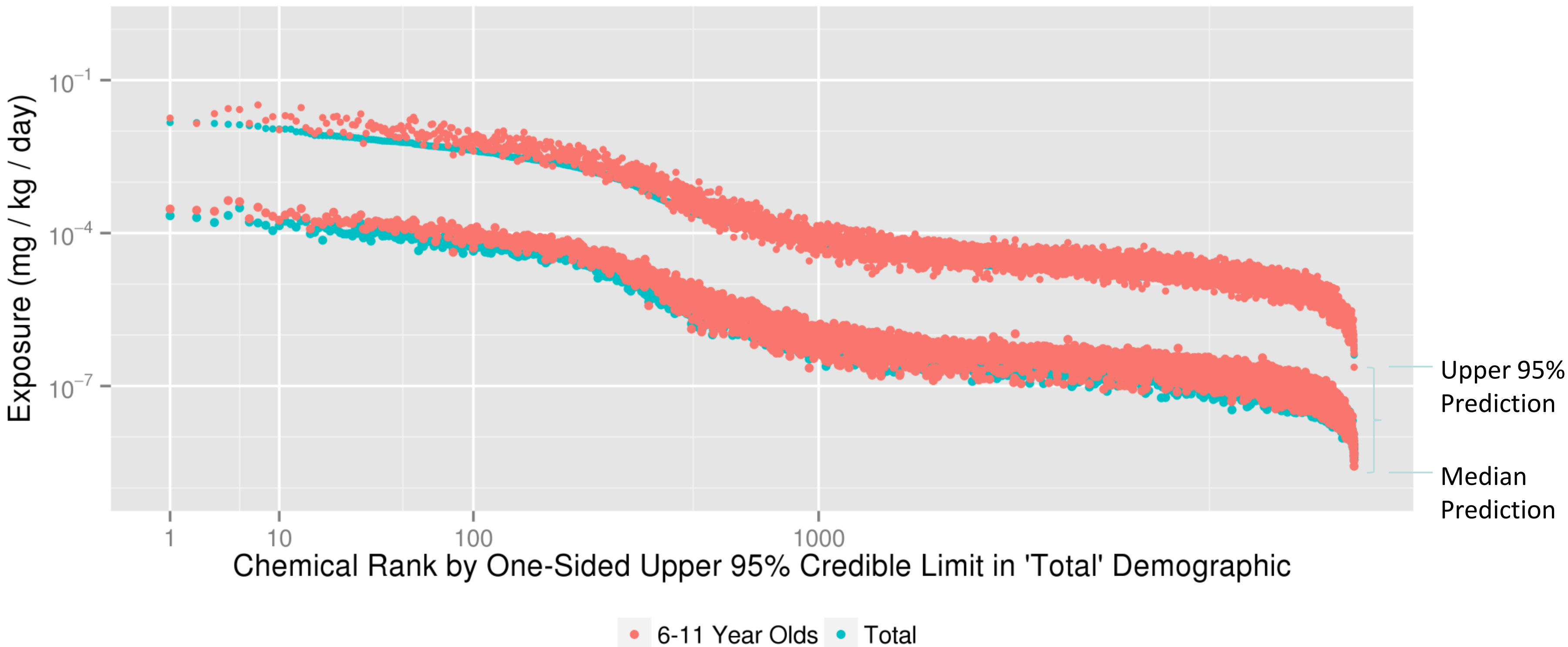
Wambaugh *et al.*, 2014

High-throughput exposure heuristics

Heuristic	Description	Number of Chemicals: Inferred NHANES Chemical Exposures (106)	Number of Chemicals: Full Chemical Library (7784)
ACToR "Consumer use & Chemical/Industrial Process use"	Chemical substances in consumer products (*e.g.*, toys, personal care products, clothes, furniture, and home-care products) that are also used in industrial manufacturing processes. Does not include food or pharmaceuticals.	**37**	683
ACToR "Chemical/Industrial Process use with no Consumer use"	Chemical substances and products in industrial manufacturing processes that are not used in consumer products. Does not include food or pharmaceuticals	**14**	282
ACToR UseDB "Pesticide Inert use"	Secondary (*i.e.*, non-active) ingredients in a pesticide which serve a purpose other than repelling pests. Pesticide use of these ingredients is known due to more stringent reporting standards for pesticide ingredients, but many of these chemicals appear to be also used in consumer products	**16**	816
ACToR "Pesticide Active use"	Active ingredients in products designed to prevent, destroy, repel, or reduce pests (*e.g.*, insect repellants, weed killers, and disinfectants).	**76**	877
TSCA IUR 2006 Total Production Volume	Sum total (kg/year) of production of the chemical from all sites that produced the chemical in quantities of 25,000 pounds or more per year. If information for a chemical is not available, it is assumed to be produced at <25,000 pounds per year.	**106**	7784

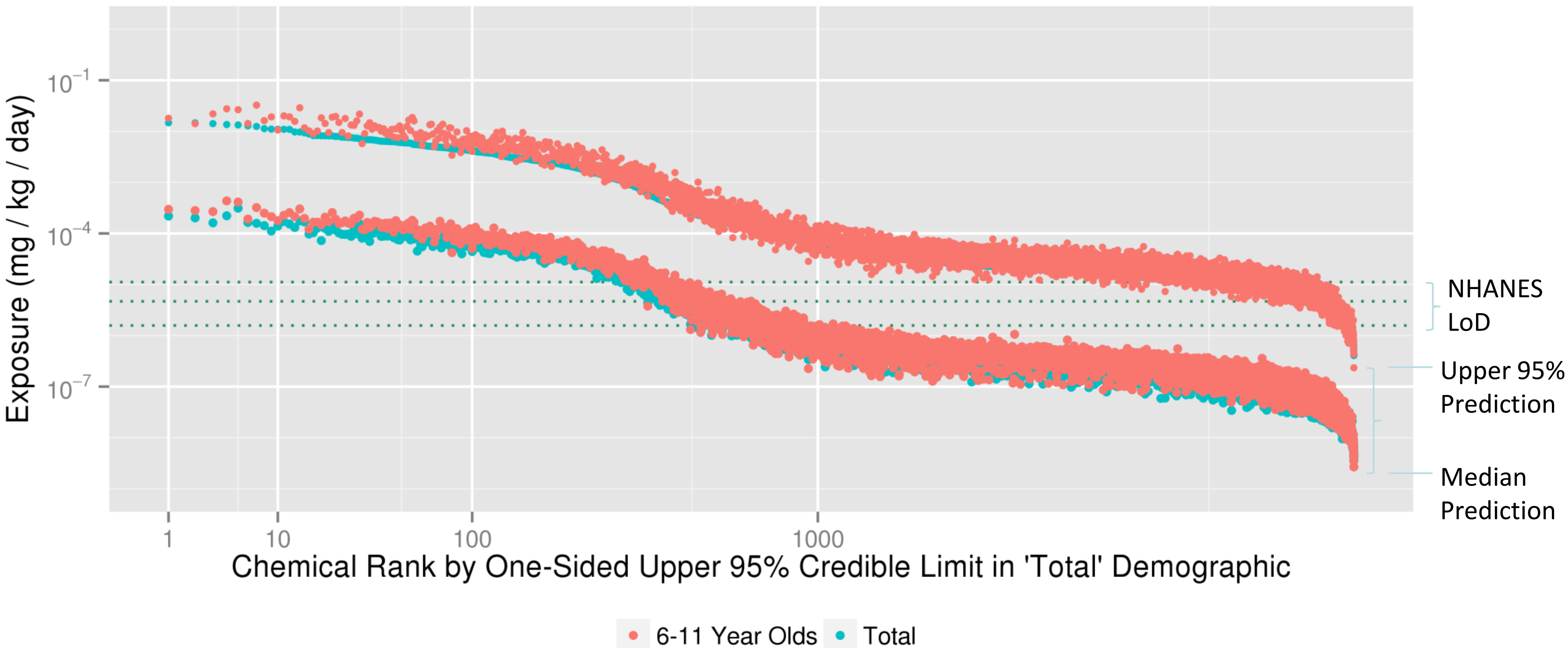
Wambaugh *et al.*, 2014

Calibrated Exposure Predictions for 7968 Chemicals



Environmental Science & Technology, *in press*

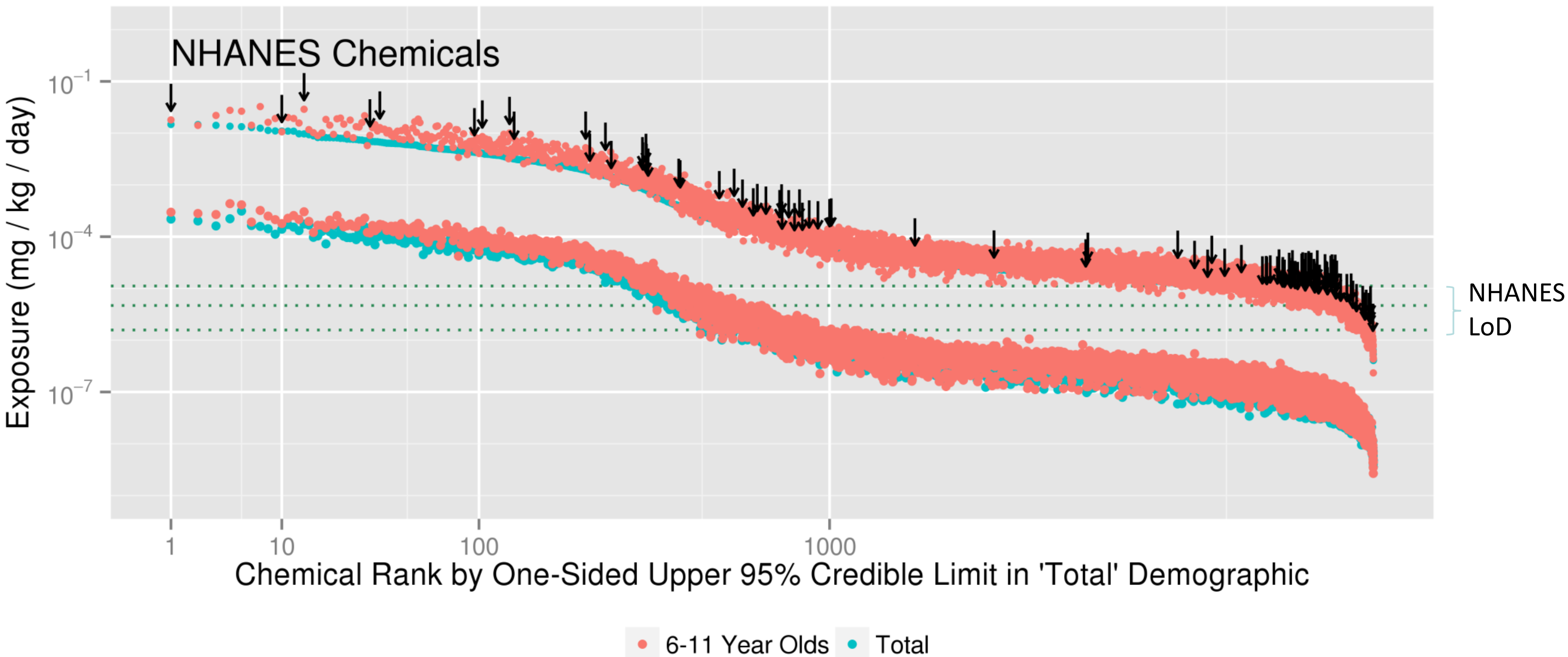
Calibrated Exposure Predictions for 7968 Chemicals



- We focus on the median and upper 95% predictions because the lower 95% is below the NHANES limits of detection (LoD)
- Dotted lines indicate 25%, median, and 75% of the LoD distribution

Environmental Science & Technology, *in press*

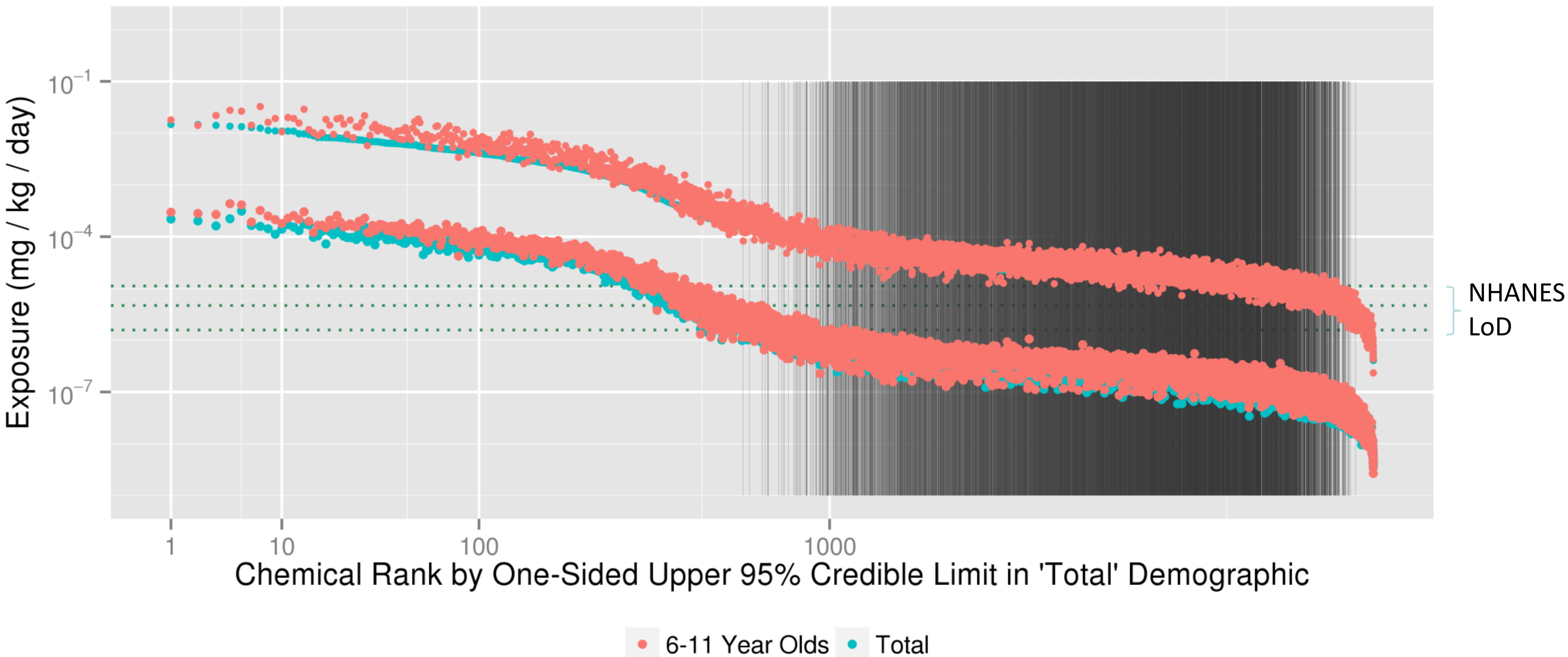
Calibrated Exposure Predictions for 7968 Chemicals



- Chemicals currently monitored by NHANES are distributed throughput the predictions
- Chemicals with the first and ninth highest 95% limit are monitored by NHANES

Environmental Science & Technology, *in press*

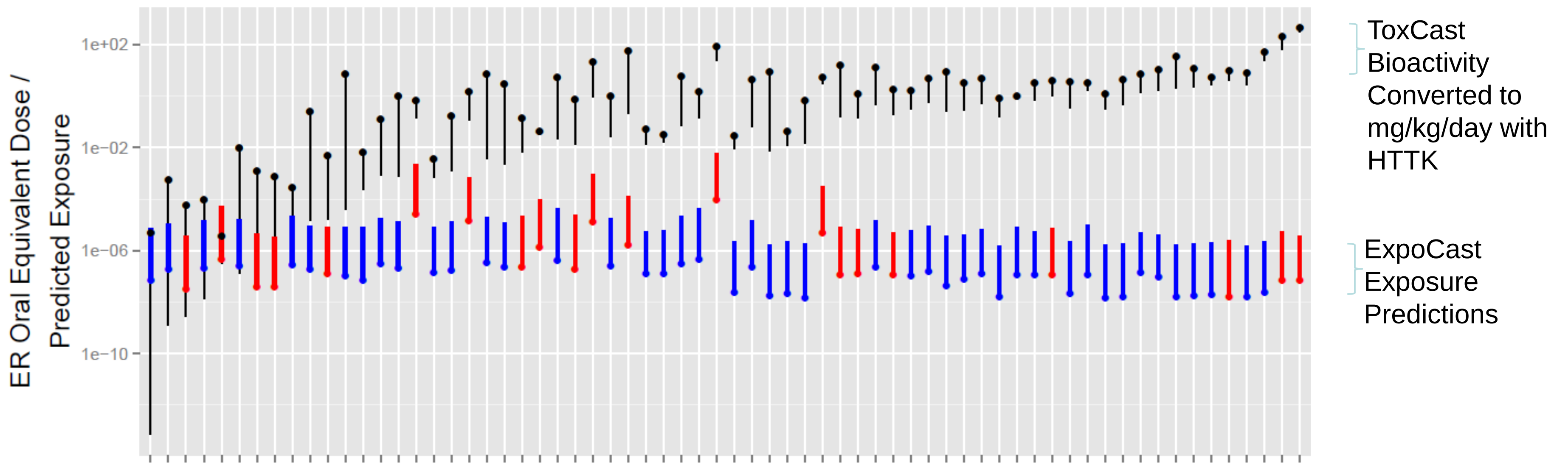
Calibrated Exposure Predictions for 7968 Chemicals



- The grey stripes indicate the 4182 chemicals with no use indicated by ACToR UseDB for any of the four use category heuristics

Environmental Science & Technology, *in press*

IBER Scientific Advisory Panel (SAP)



ToxCast Chemicals

Prioritization as in Wetmore *et al*. (2012) Bioactivity, Dosimetry, and Exposure Paper

December, 2014 Panel:
“Scientific Issues Associated with Integrated Endocrine Bioactivity and Exposure-Based Prioritization and Screening“

DOCKET NUMBER:
EPA–HQ–OPP–2014–0614

EPA United States Environmental Protection Agency

Refined Models and Better Data

Chemical to Chemical Variability of NHANES Biomonitoring

~10% Far field (Industrial) Releases
Wambaugh et al. (2013)

~50% Indoor / Consumer Use
Wambaugh et al. (2014)

Consumer product database and two new near field models

Contents lists available at ScienceDirect
Food and Chemical Toxicology
journal homepage: www
ELSEVIER

Development of a consumer produ
exposure screening and prioritizati

M.-R. Goldsmith [a,*], C.M. Grulke [a], R.D. Brook
R. Edwards [d], D.T. Chang [a], R. Tornero-Velez [a]
J. Mitchell [g], D.A. Vallero [a], L. Phillips [a], M. Ph
T.J. Buckley [a], C.C. Dary [a]

[a] United States Environmental Protection Agency (U.S. EPA), Office of Resear
[b] Student Services Contractor at U.S. EPA, RTP, NC, United States
[c] Lockheed-Martin Information Technology, RTP, NC 27711, United States
[d] North Carolina State University, 2200 Hillsborough St., Raleigh, NC 27695,
[e] Oak Ridge Institute for Science and Education Fellow, United States
[f] University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, United
[g] Biosystems and Agricultural Engineering, Michigan State University, E. Lan

ARTICLE INFO

Article history:
Received 16 October 2013
Accepted 19 December 2013
Available online 27 December 2013

Keywords:
Chemical exposure
Consumer products
Ingredients
Product formulation
Near field exposure
Exposure prioritization

ABSTRAC

Consumer produ
able on the chen
ent. To address
Material Safety
sents 1797 uniq
uct "use categor
discuss ways in
formulations for
selection for mo
uitous exposure
and across mult
fied. Our databa
predictive scree
risk.

Environmental Science & Technology
Article
pubs.acs.org/est

Model for Screening-Level Assessm
Exposure to Neutral Organic Che

Xianming Zhang,[*,†,§] Jon A. Arnot,[*,†,‡] and Fran

[†] Department of Physical and Environmental Sciences, Univers
[‡] ARC Arnot Research and Consulting, Toronto, Ontario M4M

Supporting Information

ABSTRACT: Screening organic chemicals for hazard and ri
health requires near-field human exposure models that ca
parametrized with available data. The integration of a mod
exposure, uptake, and bioaccumulation into an indoor mass b
provides a quantitative framework linking emissions in ind
ments with human intake rates (*iRs*), intake fractions (*iFs*)
state concentrations in humans (C) through consideratio
permeation, inhalation, and nondietary ingestion exposu
Parameterized based on representative indoor and a
characteristics, the model is applied here to 40 chemicals
in the context of human exposure assessment. Intake fraction
concentrations (C_U) calculated with the model based on a u
rate to air for these 40 chemicals span 2 and 5 orders o
respectively. Differences in priority ranking based on either *i*
elimination processes within the human body. The model
representative of many in-use chemicals to show how the
chemical properties and to illustrate the capacity of the m

Environmental Science & Technology
Article
pubs.acs.org/est

SHEDS-HT: An Integrated Probabilistic Exposure Model for Prioritizing Exposures to Chemicals with Near-Field and Dietary Sources

Kristin K. Isaacs,[*,†] W. Graham Glen,[‡] Peter Egeghy,[†] Michael-Rock Goldsmith,[§,○] Luther Smith,[‡] Daniel Vallero,[†] Raina Brooks,[∥] Christopher M. Grulke,[⊥,○] and Halûk Özkaynak[†]

[†] U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory, 109 T.W. Alexander Drive, Research Triangle Park, North Carolina 27709, United States
[‡] Alion Science and Technology, 1000 Park Forty Plaza Suite 200, Durham, North Carolina 27713, United States
[§] Chemical Computing Group, Suite 910, 1010 Sherbrooke Street West, Montreal, QC H3A 2R7, Canada
[∥] Student Services Contractor at U.S. Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park, North Carolina 27709, United States
[⊥] Lockheed Martin, 109 T.W. Alexander Drive, Research Triangle Park, North Carolina 27709, United States

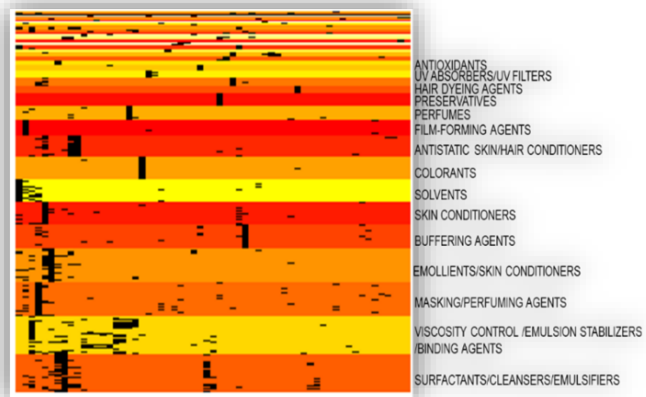
Supporting Information

ABSTRACT: United States Environmental Protection Agency

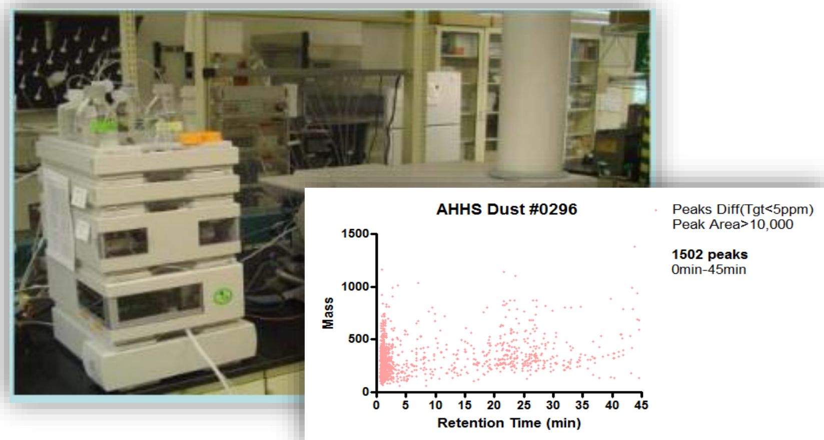
Consumer Product Data

New Exposure Related Data

United States
Environmental Protection
Agency



New Chemical Use Information



New Monitoring Data

New data on chemicals within and emission from consumer products

nielsen

Consumer Product Use Information

New data on physico-chemical properties

CSS Rapid Exposure and Dosimetry

Chemical Use Information for >30,000 Chemicals

- Chemical-Product Categories (CPcat) database maps many different types of use information and ontologies onto each other
- Includes CPCPdb (Goldsmith, et al., 2014) with information on >2000 products from major retailors
- Largest single database has coarsest information: ACToR UseDB

Table: Hits per use category for a given chemical

CASRN	Category 1	Category 2	…	Category 12
65277-42-1	0	10	…	1
50-41-9	31	7	…	3
…	…	…	…	…

Binary matrix

CASRN	Category 1	Category 2	…	Category 12
65277-42-1	0	1	…	0
50-41-9	1	1	…	0
…	…	…	…	…

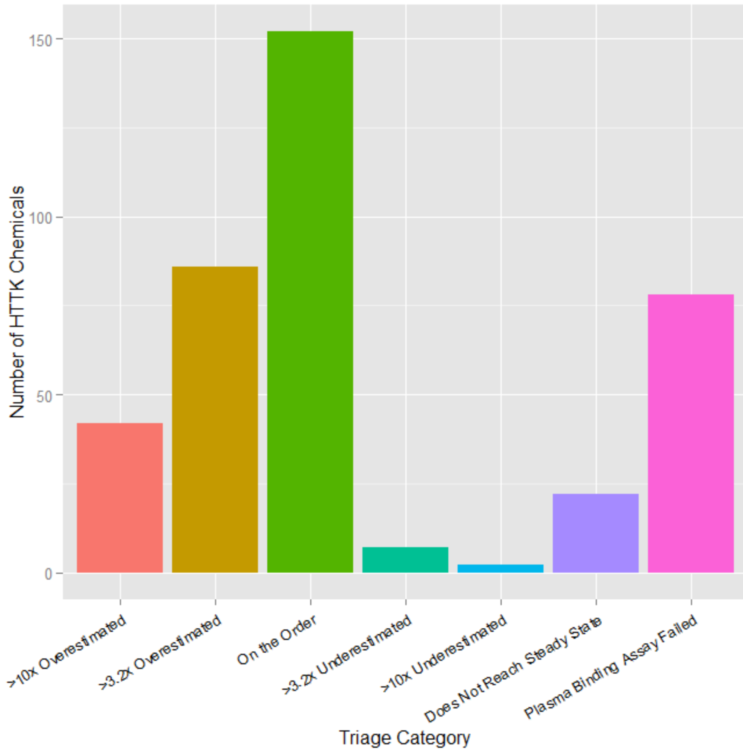
12 Chemical Use Categories
Antimicrobials
Chemical Industrial Process
Consumer
Dyes and Colorants
Fertilizers
Food Additive
Fragrances
Herbicides
Personal Care Products
Pesticides
Petrochemicals
Other

http://actor.epa.gov/cpcat/

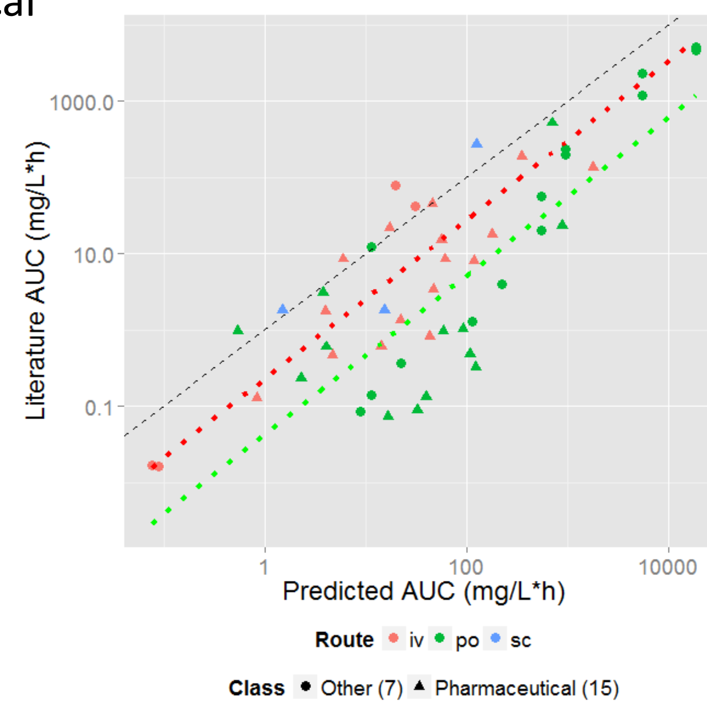
Dionisio et al. *(2015)*

United States Environmental Protection Agency

High Through Toxicokinetics



- We now have high throughput *in vitro* data allowing toxicokinetic (TK) predictions for ~500 chemicals – ToxCast is collecting more
- We can use *in vitro* data, physico-chemical properties, high throughput physiologically-based TK models, and transporter QSARs to decide whether HTTK methods are appropriate for a given chemical



- All tools and data publically available on CRAN (R package "httk")
- We are also developing an *in vivo* data set for better evaluation of methods for environmental chemicals

STAR Grantees New Methods in 21st Century Exposure Science

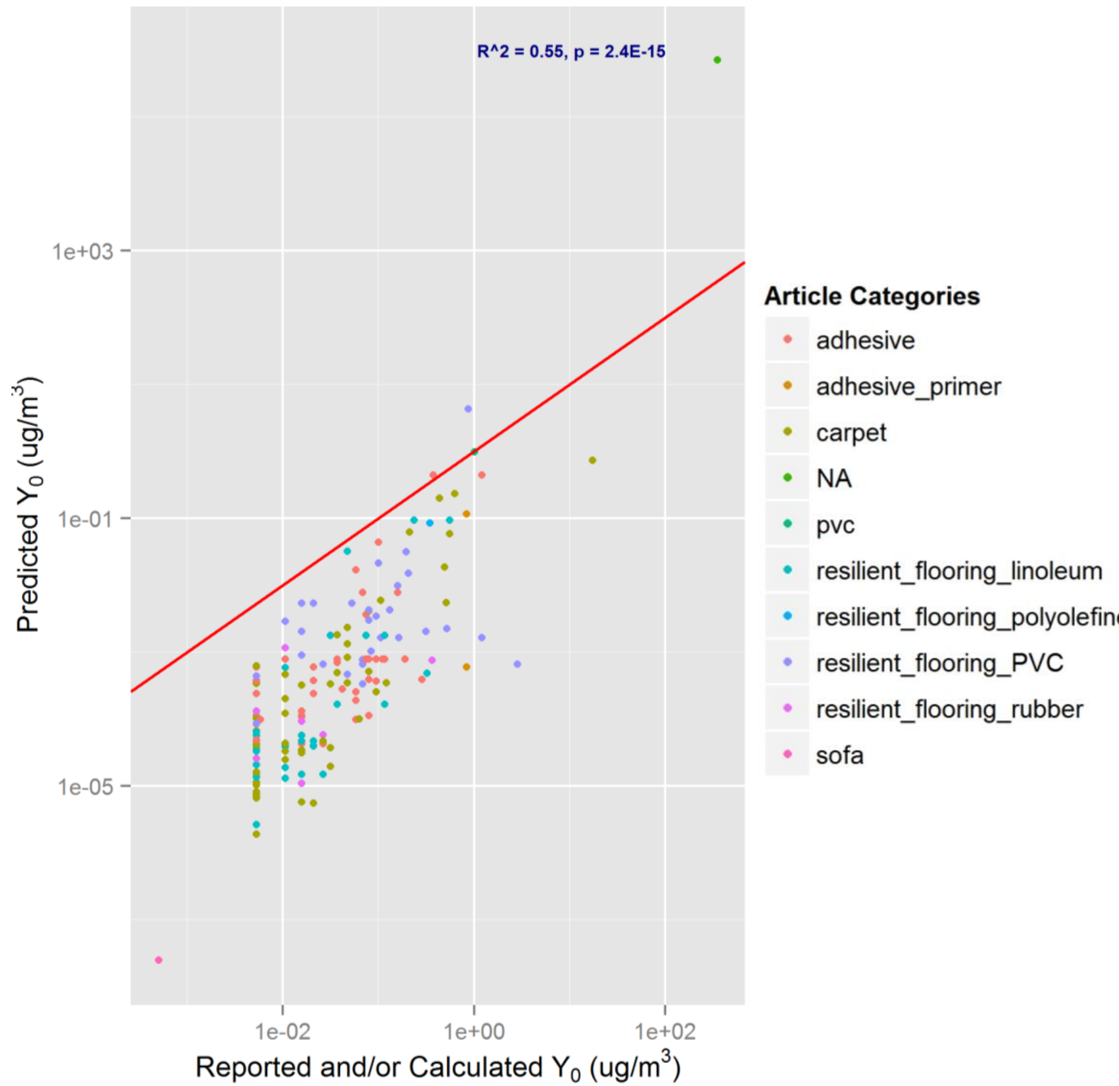
Principal Investigator	Institution	Title of Application
Bennett, Deborah	University of California, Davis	Tracking Semivolatile Organic Compounds (SVOCs) Indoors: Merging Models and Field Sampling to Assess Concentrations, Emissions, and Exposures
Stapleton, Heather	Duke University	Residential Exposure of Young Children to SVOCs
Woodruff, Tracey	University of California, San Francisco	A Non-targeted method for measuring multiple chemical exposures among a demographically diverse population of pregnant women in Northern California
Fan, Xudong	University of Michigan	3-dimensional micro-gas chromatography device for rapid and sensitive indoor air exposure assessment
Little, John	Virginia Tech	Rapid methods to estimate exposure to SVOCs in indoor environments

Gas-Phase Concentration Model

- 73 total chemicals in model including SVOCs[1] reported from Wilke *et al*. (2004)

- 4 chemicals reported from Little *et al.* (2012)

- 1 main physicochemical property that model data (VP). Other predictors include formulation descriptors.



Acronyms:
- SVOCs = Semivolatile Organic Compounds
- FRs = Flame Retardants
- VP = Vapor Pressure
- $Y_{\circ}$ = Gas-phase concentration

Slide from Chantel Nicolas

ExpoCast Contracts Awarded in December 2014

Exposure Screening Tools for Accelerated Chemical Prioritization (ExpoCast)

- Two awardees:
  Battelle Memorial Institute (Columbus, OH) and
  Southwest Research Institute (San Antonio, TX)

- The EPA is interested in building models to quantitatively predict potential exposure for thousands of chemicals in commerce. Results will be used in the ExpoCast project to evaluate, calibrate and reduce uncertainty in exposure model predictions and for prioritizing compounds for more in-depth testing and risk assessment. To support computational models three kinds of exposure measurement data are required:

  (1) key physical-chemical properties
  (2) chemical emissions from consumer products used indoors
  (3) chemical occurrence in products, environmental, and biological media

United States Environmental Protection Agency

ExpoCast Contracts

- We refine exposure predictions and characterize uncertainty in chemical exposures by examining the predictive ability of models and the coverage (or lack thereof) of critical pathways

- Data goals for ExpoCast contracts:

  (1) Key physical-chemical properties:
    - Model predictions depend on physico-chemical properties, for some classes of relevant chemicals these properties have never been measured
    - We wish to augment existing predictive quantitative structure-activity relationships (QSAR) with measurements for new classes

ExpoCast Contracts

- We refine exposure predictions and characterize uncertainty in chemical exposures by examining the predictive ability of models and the coverage (or lack thereof) of critical pathways

- Data goals for ExpoCast contracts:

  (2) Chemical emissions from consumer products used indoors:
    - If we are lucky enough to figure out the chemical composition of a product, we need data to characterize emission of the chemical from the product
    - Our goal is to build sufficiently diverse library (representative chemical-product combinations) to allow QSAR estimates of emission

ExpoCast Contracts

- We refine exposure predictions and characterize uncertainty in chemical exposures by examining the predictive ability of models and the coverage (or lack thereof) of critical pathways

- Data goals for ExpoCast contracts:

  (3) Chemical occurrence in product, environmental, and biological media

  - Product deformulation – linking chemicals to products helps determine exposure pathways
    - There are many chemicals (e.g. packaging materials) that are not deliberately added by final product manufacturer
  - Enhanced monitoring data
    - We need to better characterize exposures both to more diverse human populations and to more diverse chemistries (e.g., non-targeted screening)

Interpreting the Exposome Requires Insight at All Levels of Organization

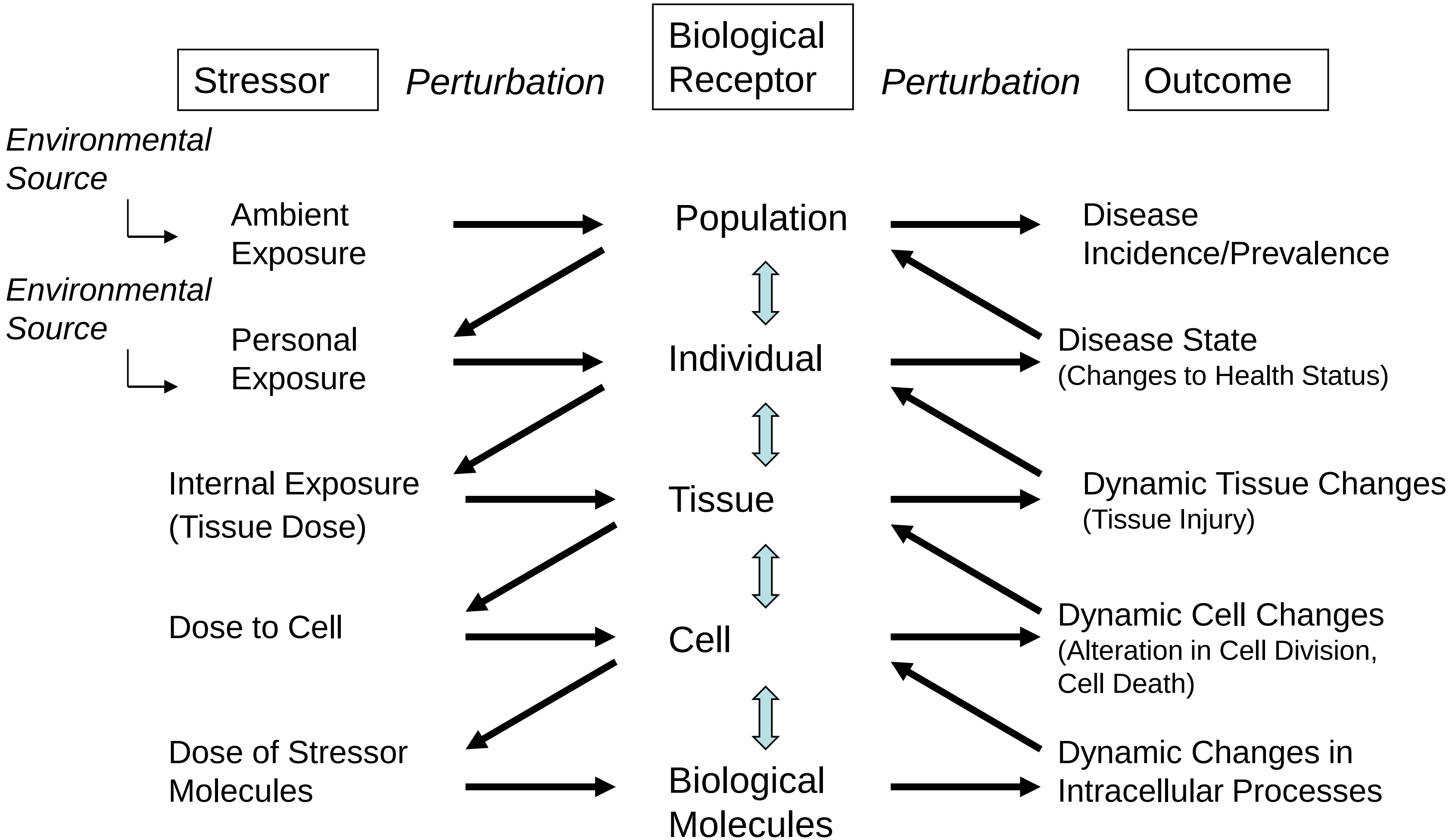


Figure modified from Cohen Hubal et al. (2008)

Conclusions

- Interpreting the Exposome requires insight at all levels of biological and social organization
- There are low levels of thousands of chemicals present in the metabolome, relating these to exposures and health effects is an important unsolved problem
- The exposure pathway is the actual interaction of the receptor and media, and this event is often confounded by various sources of uncertainty
- Can use a combination of forward modeling and reverse inference from biomarkers of exposure to examine plausible exposures – there is a great need to carefully consider predictive performance with statistical analysis

EPA United States Environmental Protection Agency

Chemical Safety for Sustainability (CSS) Rapid Exposure and Dosimetry (RED) Project

NCCT
Chris Grulke
Richard Judson
Chantel Nicolas*
Robert Pearce*
James Rabinowitz
Ann Richard
Caroline Ring*
Woody Setzer
Rusty Thomas
John Wambaugh

NRMRL
Yirui Liang
Xiaoyu Liu*

NHEERL
Jane Ellen Simmons
Marina Evans
Mike Hughes

NERL
Craig Barber
Derya Biryol*
Dalizza Colon
Kathie Dionisio*
Marc Duchatelier*
Peter Egeghy
Kim Gaetz
Brandall Ingle*
Kristin Isaacs
Katherine Phillips*
Paul Price
Julia Rager*
Mark Strynar
Jon Sobus
Mike Tornero-Velez
Elin Ulrich
Dan Vallero

*Trainees

Collaborators

Arnot Research and Consulting
Jon Arnot
Chemical Computing Group
Rocky Goldsmith
Environmental Protection Agency
Alicia Frame
Hamner Institutes
Barbara Wetmore
Cory Strope
Indiana University
James Sluka
Michigan State University
Jade Mitchell
National Institute for Environmental Health Sciences (NIEHS)
Mike Devito
Nisha Sipes
Kyla Taylor
Kristina Thayer
Netherlands Organisation for Applied Scientific Research (TNO)
Sieto Bosgra
North Carolina State University
Anran Wang*
Research Triangle Institute
Timothy Fennell
Silent Spring Institute
Robin Dodson
University of California, Davis
Deborah Bennett
University of Michigan
Olivier Jolliet
University of North Carolina, Chapel Hill
Alexander Sedykh*
Alex Tropsha

References

- Dionisio, Kathie L., et al. “Exploring Consumer Exposure Pathways and Patterns of Use for Chemicals in the Environment.” Toxicology Reports (2015)
- Egeghy, Peter P., et al. "The exposure data landscape for manufactured chemicals." Science of the Total Environment 414: 159-166 (2012)
- Goldsmith, M-R., et al. "Development of a consumer product ingredient database for chemical exposure screening and prioritization." *Food and chemical toxicology* 65 (2014): 269-279.
- Isaacs, Kristin K., et al. "SHEDS-HT: An Integrated Probabilistic Exposure Model for Prioritizing Exposures to Chemicals with Near-Field and Dietary Sources.” Environmental Science and Technology 48.21 (2014): 12750-12759.
- Little, John C., et al. "Rapid methods to estimate potential exposure to semivolatile organic compounds in the indoor environment." *Environmental science & technology* 46.20 (2012): 11171-11178.
- National Research Council. "Exposure Science in the 21st Century: A Vision and a Strategy." (2012).
- Park, Youngja, H., et al. “High-performance metabolic profiling of plasma from seven mammalian species for simultaneous environmental chemical surveillance and bioeffect monitoring.” Toxicology 295:47-55 (2012)
- Pilkey, Orrin H., and Linda Pilkey-Jarvis. *Useless arithmetic: why environmental scientists can't predict the future*. Columbia University Press, 2007.
- Rosenbaum, Ralph K., et al. “USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment” International Journal of Life Cycle Assessment 13: 532-546 (2006)
- Silver, Nate. *The signal and the noise: Why so many predictions fail-but some don't*. Penguin, 2012.
- Wallace et al., “The TEAM Study: Personal exposures to toxic substances in air, drinking water, and breath of 400 residents of New Jersey, North Carolina, and North Dakota .” Environmental Research 43: 209-307 (1987)
- Wambaugh, John F., et al. "High-throughput models for exposure-based chemical prioritization in the ExpoCast project." *Environmental science & technology* 47.15 (2013): 8479-848.
- Wambaugh, John F., et al. "High Throughput Heuristics for Prioritizing Human Exposure to Environmental Chemicals." *Environmental science & technology* (2014).
- Wild, Christopher P., “Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology.” Cancer Epidemiology, Biomarkers, and Prevention, 14: 2487 (2005)
- Wilke, Olaf, Oliver Jann, and Doris Brödner. "VOC-and SVOC-emissions from adhesives, floor coverings and complete floor structures." *Indoor Air* 14.s8 (2004): 98-107.